ORGANS LEARNING USING CNN

Sandhya M Y, MCA Student, PES Institute Of Technology And Management,
Shivamogga, Karnataka, India.

Ms. Kavya H V, Assistant Professor Dept. of MCA, PES Institute Of Technology
And Management, Shivamogga, Karnataka, India.

ABSTRACT

The limited visual perception hinders the movement and independence of the visually impaired. This research presents a Convolutional Neural Networks (CNNs) based system to automatically identify organs in medical images, with its application to navigation and learning through real-time objects recognition. The system detects human organs (brain, heart, liver, kidney, eye, hand and leg) through the use of weakly supervised object detection techniques. The model achieves an accuracy of 95% using transfer learning in conjunction with the InceptionV3 architecture. It is accessible using desktop and mobile platforms with easy to use interface with comprehensive textual and video information to provide highest level of accessibility. The target is to resolve the inequities in instruction and assistive applied sciences by expeditiously locating organs by applying deep intelligence.

Keywords : *Convolutional Neural Network (CNN), Organ Detection, Image Classification, Transfer Learning, Weakly Supervised Learning, InceptionV3, Object Recognition, Visually Impaired Assistance.*

I. INTRODUCTION

People with sight problems would tend to develop problems with the recognition and understanding of sight items in the environment. This problem can be solved using object recognition since it has the capacity to identify and present the contents of the pictures in terms of sensations. The main objective of traditional object detection methods is to detect a particular object that is not in motion; in the proposed study, the three Convolutional neural networks (CNNs) which are a form of medical learning method are utilized to categorize images with reference to the deep anatomy of the organs.

This project will come up with a system of detecting incidences in real time, which might independently cluster pictures of prominent human organs. Data on the objects found are presented in the form of text and videos, which enables easy understanding and memorizing. The system has been lightened, easily stretchable and efficient both in the desktop and mobile environment.

II. LITERATURE SURVEY

On the one hand, the convolution model and on the other hand non-linear relationships model are the two basic models which comprise the deep neural networks. The thing is regarded in the models as a stratified object, which is represented as a structure of simpler objects. Over time, various architectures and algorithms have been introduced to implement deep learning, such as belief networks, stacked networks, and gated recurrent units, among others. They had achieved the invention of the initial convolutional neural network (CNN). Convolutional neural networks may be applied to the recognition of handwriting and image processing. Object detection involves determining both the coordinates and the categories of specific objects within an image, regardless of their position. In this project, the focus is on utilizing either the Faster R-CNN architecture or the YOLOv3 framework for this task. Faster R-CNN Region Proposal Network entails the creation of regions, and in two Faster R-CNN, modes, the identification of objects. The first one is the approach, which implies the suggestion of using suggested regions. The author of Faster RCNN used a convolutional network wherein the 16-layer structure was set to the aim of attaining the detection and classification accuracies over various datasets. Another protocol given by Kumar, A. et al. [1] is the buyer-seller watermarking to offer security and privacy to the trade between the communicating parties. YOLO v3 is an acronym, You Only Look Once. It is real time object detector where it detects the objects using learnt features through the aids of deep convolutional neural network. YOLOv3 consists of 75 convolutional layers, incorporating upsampling layers and skip connections to enable processing of the entire image through a single neural network. It generates specific regions of interest, producing corresponding bounding boxes along with their confidence scores. A key feature of YOLOv3 is its ability to perform detections at three different scales.

III. SYSTEM ANALYSIS AND DESIGN

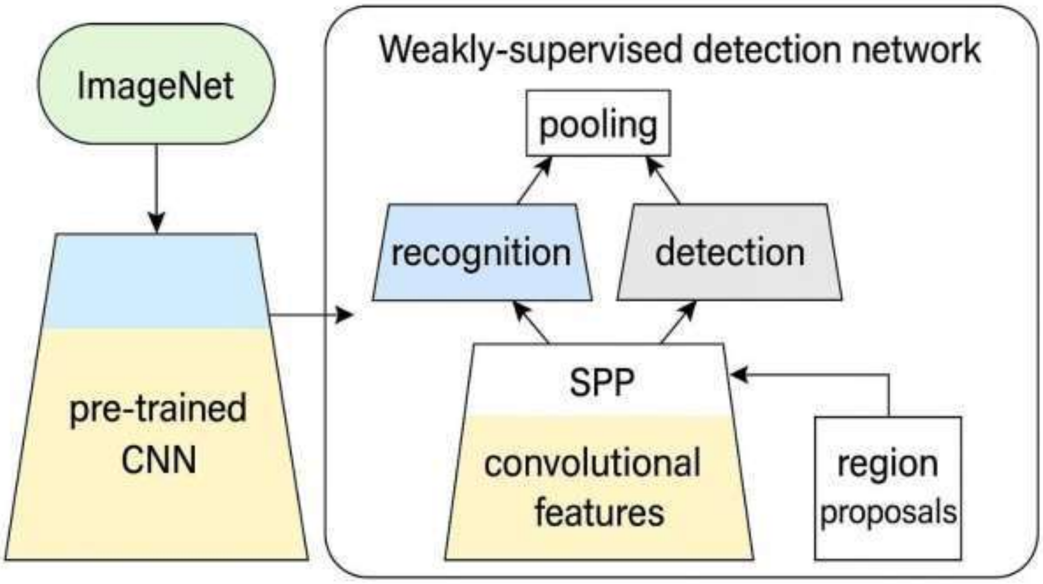


Fig 3.1: System architecture diagram

Some of the advantages of the model are the tiny size, energy-efficiency, and the possibility to prove low error rates as a result of being fully-convolutional and having only one forward pass. Figure 1 is the summary of the process in the class of this object detection model. The CNN which we were using is referred to as SqueezeDet. SqueezeDet is the whole convolutional network which strives to carry out object detection. It is an extension of SqueezeDet architecture, where feature maps of images are generated through the utilization of CNN. Next to this, it uses another convolution layer to extract the coordinates of bounding boxes, the confidence levels and probability of the classes. Multi target loss is employed at some point to estimate the final loss in the training part and Non- Maximum Suppression (NMS) applied in attempts to decrease the overlapping bounding boxes such that the final result is the final detection in the challenge section where testing takes place.

IV. HARDWARE AND SOFTWARE REQUIREMENTS

Developing the organ detection system using Convolutional Neural Networks (CNNs) requires meeting certain hardware and software specifications. On the hardware side, a minimum of a Pentium IV or newer processor and at least 2 GB of RAM is recommended to ensure efficient and faster processing.

Also, it should have a hard disk drive of at least 80GB to accommodate the dataset, application tools and other components.

As far as software is concerned, the implementation will use Python and its primary programming language given that it significantly supports machine learning and image processing libraries. The user interface, which manifests itself in a web-page, is implemented in HTML and JavaScript, whereas image processing and manipulation involving management and preprocessing is performed in OpenCV. Finally, MySQL is the backend database that is providing the convenience of managing user authentication and application details successfully.

V. Methodology

It is anticipated to identify and classify human organs in real time applications with the implementation of deep learning techniques such as the Convolution Neural Networks (CNNs). The overall paradigm is transfer learning, weakly supervised learning, and the detection of objects in real-time. The system begins by obtaining a comprehensive set of images for various organs, including the brain, eye, heart, kidney, liver, hand, and leg. These images may consist of medical scans such as CT and MRI, as well as online, virtual, and real-time captured photographs. Preprocessing on images is done before training, nomenclature such as resizing, normalization, and augmentation, i.e., rotations, flip and zoom commands are applied as a training model makes it stronger and therefore the model that can generalize better. The system employs transfer learning to reduce the time and computational cost required for training by utilizing the InceptionV3 model, which has been pre-trained on the ImageNet dataset. In this approach, the lower layers of the InceptionV3 network responsible for learning fundamental features are kept frozen to preserve their learned parameters. Additional dense layers are then added and trained to classify the specific organ parts targeted in this project. The major process in the object detection process is the processing of the images into the CNN where successive convolution layers extract the hierarchical features such as the textures, shapes and edges. The features becoming obtained are compressed and this process is introduced to the fully connected layers, which keeps track of the entire classes of the organs through a softmax activation layer.

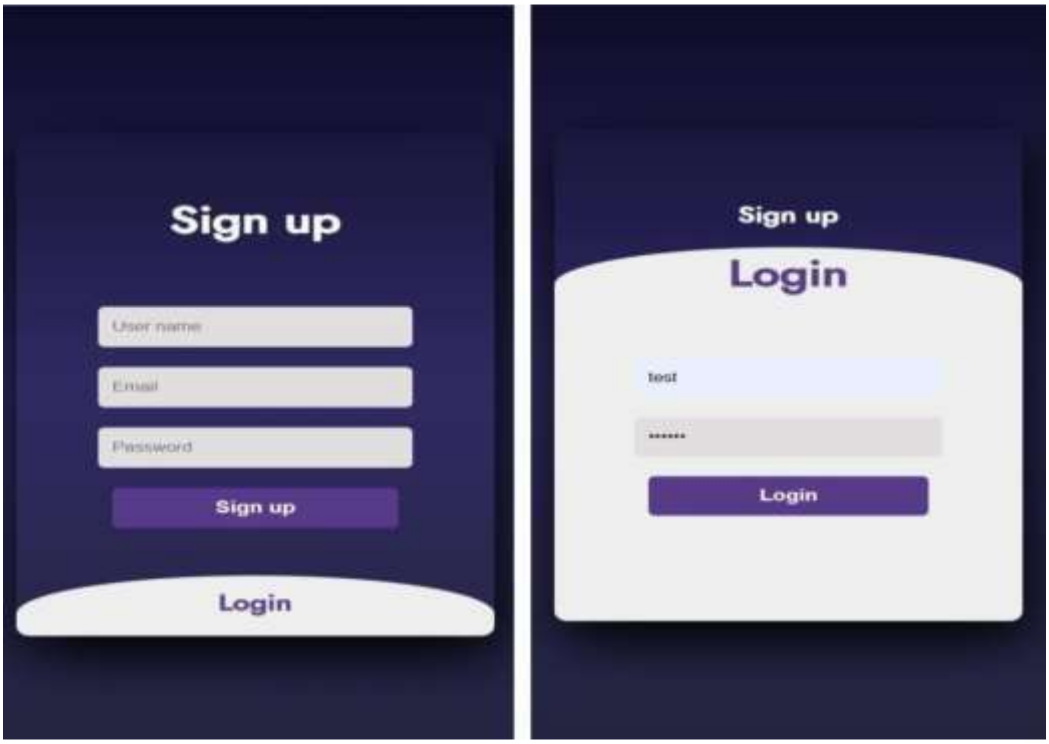
This project utilizes the SqueezeDet and YOLOv3 architectures for object identification. These models are well-suited for rapid object recognition in images due to their ability to perform multiple tasks simultaneously specifically, predicting object locations through bounding boxes and classifying them within a single process. Their efficiency enables the system to operate effectively in real-time applications. Also, the system incorporates Weakly Supervised Detection (WSD), as a result, the CNN model can attain the knowledge on the classification of objects which requires minimal form of labeling as only images are labeled but not bounding boxes.

VI. Implementation

The coding in the project utilizes Python, Tensor Flow/Keras, and OpenCV, whereas Flask was used to develop a backend that is compatible with the web. The system's user interface allows users to log in by entering a password, with credentials securely stored in a MySQL database for registration or access to additional features. Image inputs can be provided in several ways: by selecting files already saved on the device, capturing new photos using a connected camera, or uploading images from the internet.

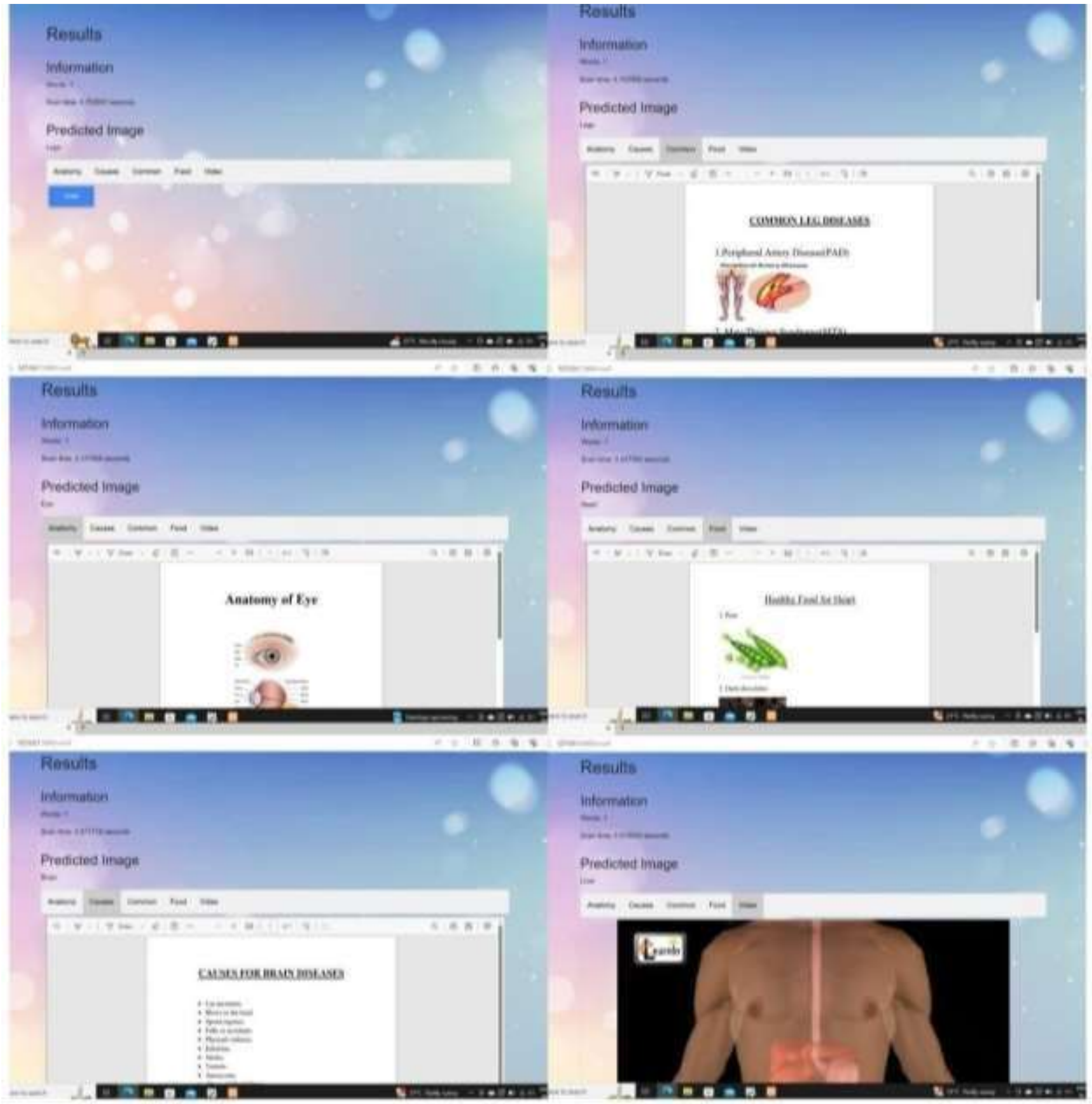
After the uploading, an image is passed through a trained CNN network that recognizes the object in question and derives the relevant medical facts. The results can be viewed on a simple to use user interface where one can inspect the anatomy, common diseases, causative factors, dietary advice, and even training. The results of the study are described in a convenient interface that carries informational movies, an up close anatomy experience, the most common ailments and their causes, and dietary guidelines. Such a trained model could be saved as model_inception.h5 in order to avoid any difficulties in using it in desktop application and on a mobile device. Upon taking a trail, the system was found to be precise when it came to storing individuals photographs, the space image, and a real-time photograph at the rate of 95, 70, 60 respectively. The results of processing each photo were on average 2-4 seconds. The provided implementation demonstrates the scalable, cost effective, and efficient method of organ identification and training and can be practical to apply in healthcare training and assistance, as well as assisting patients with eye. The provided implementation presents a low-cost, efficient, and scalable method of organ detection and education, and it is viable to utilize in healthcare education and assistance, as well as help the patients with low vision.

VII. Results and Conclusion



By Just entering the IP address in Address bar we enter the sign-in page. In this case, we place our username and password to sign up. In the situation, where we are already registered, we just push the icon of the log in and write our user name and password to pass to the next page.

On this page, we give you a choice of selecting image either by uploading one to file on your device, or by searching it with the help of Google. While using this feature we can also able to take pictures using our mobile phone. After choosing a picture, you should press the predict image button and on the bottom of the page you will find the picture of which organ is chosen.



We will know the organ in this section and in this section we shall have five options namely Anatomy, Causes, Common Diseases, Food, and Video. Upon clicking the anatomy, you will obtain text facts of that organ and the name of the organ and the functionality of the organ. You can select the causes of your choice, and you will receive the information about those factors which predispose diseases of the specific organ. Experience gained by clicking on common diseases will get you the information on common states of the organ. Just by clicking the food it will inform you what are the healthy diet food option which will be favourable to the respective organ. Finally, the choice of the video will provide any types of information in a video form.

Conclusion

Object detection by deep learning was the research hotspot in the recent years. This 3233 1616 begins with generic object detection pipelines that has baseline architecture of other related tasks. In general, common applications of this work include object detection, face recognition, and pedestrian identification. The authors accomplished pedestrian detection by combining two key components: deep learning-based object detection and OpenCV for efficient, multi-threaded video streaming. However, factors such as camera sensor limitations, noise, and varying lighting conditions can impact accuracy, often leading to object identification issues. The resulting system is a deep learning object detector capable of achieving an average processing speed of 6–8 FPS. For this project, a CNN-based object detection model was trained on a desktop platform and later deployed to a mobile device for practical implementation. As a baseline, we have a running and android app that runs our CNN model trained by Tensorflow offline. The model size is 8 MegaBytes and achieved testing accuracy is 95%.

VIII. Future Improvements

The concept is to know how to identify discriminative local patches automatically and how to identify local feature information in recognizing parts of different bodies. Importantly, the specified approach does not entail any manual labeling to mark such local patches, and it is highly scalable and massive labeling operations are not needed. In the future, we can supplement some additional details and enhance the approach, concerning the language. As this builds up, it can also serve on higher education and clinical practice etc.

References

[1] Petrov, Yordan. Improving object detection by exploiting semantic relations between objects. MS thesis. Universitat Politècnica de Catalunya, 2017

[2] Nikouei, Seyed Yahya, et al. "Intelligent Surveillance as an Edge Network Service: from Harr-Cascade, SVM to a Lightweight CNN." arXiv preprint arXiv:1805.00331 (2018).

[3] R. Girshick. Fast r-cnn. In Proc. ICCV, 2015.

[4] Howard, Andrew G., et al. "Mobilenets: Efficient convolutional neural networks for mobile vision applications." arXiv preprint arXiv:1704.04861 (2017).

[5] Kong, Tao, et al. "Ron: Reverse connection with objectness prior networks for object detection." 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).IEEE, 2017.

[6] Liu, Wei, et al. "Ssd: Single shot multibox detector." European conference on computer vision.Springer, Cham, 2016. Veiga, Francisco José Lopes. "Image Processing for Detection of Vehicles In Motion." (2018)

[7] Huaizheng Zhang, Han Hu, GuanyuGao, Yonggang Wen, Kyle Guan, "Deepqoe: A Unified Framework for Learning to Predict Video QoE", Multimedia and Expo (ICME) 2018 IEEE International Conference on, pp. 1-6, 2018. Shijian Tang and Ye Yuan,"Object Detection based on Conventional Neural Network".

[8] Shilpisingh et al" An Analytic approach for 3D Shape descriptor for face recognition", International Journal of Electrical, Electronics, Computer Science & Engineering (IJEECSE), Special Issue - ICSCAAIT-2018| E-ISSN: 2348-2273 | P-ISSN: 2454-1222,pp-138- 140.

[9] M. Jaderberg, K. Simonyan, A. Zisserman, and K. Kavukcuoglu.

Spatial transformer networks. 2015

[10] K. Simonyan and A. Zisserman. Very deep convolutional networks for large- scale image recognition. In International Conference on Learning Representations, 2015.

[11] K. Chatfield, K. Simonyan, A. Vedaldi, and A. Zisserman. Return of the devil in the details: Delving deep into convolutional nets. In Proc. BMVC., 2014.

[12] M. P. Kumar, B. Packer, and D. Koller. Self-paced learningfor latent variable models. In NIPS , 2010

[13] B. Hariharan, P. Arbelaez, R. Girshick, and J. Malik. Simultaneous detection and segmentation. In Proc. ECCV,, 2014.

[14] H. O. Song, R. Girshick, S. Jegelka, J. Mairal, Z. Harchaoui, and T. Darrell. On learning to localize objects with minimal supervision. In Proc. ICML,, 2014.

[15] R. B. Girshick, J. Donahue, T. Darrell, and J. Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proc. CVPR, 2014